FREQUENCY OF OCCURRENCE OF NOVEL MILK PROTEIN VARIANTS IN A NEW ZEALAND DAIRY CATTLE STUDY POPULATION

A thesis presented in partial fulfilment of the requirements
for the degree of
Master of Science in Biochemistry,
Massey University,
New Zealand

Richard Gordon Burr

August, 1996

ABSTRACT

Since the discovery of genetic polymorphism within milk protein genes, a considerable volume of research has been published relating milk protein genetic variants and milk production properties. Polymorphism of milk proteins can result in two effects:

(a) changes in the biological and physico-chemical properties of systems containing the variant protein,

(b) changes in the synthesis level of variant proteins.

As a result several studies of milk protein variants have identified phenotypes which may be commercially advantageous for specific products.

Currently employed methods to determine milk protein phenotypes are generally limited to electrophoretic techniques. The gel electrophoretic techniques commonly used are able to detect most milk protein variants that differ by their net electrical charge. However single amino acid substitutions that result in a change in net charge account for only 25% of the possible substitutions that could occur. The remaining 75% of potential variants are the result of a neutral residue substituted by another neutral residue - a 'silent' variant. Thus it is likely that some substitutions, and hence genetic variants have gone undetected in the past.

The purpose of this study was to develop new methods for determining the phenotype of milk proteins, and to determine the frequency of occurrence of silent or other novel variants in a New Zealand dairy cattle study population.

Polyacrylamide gel electrophoresis (PAGE), free zone capillary electrophoresis (CE), peptide mapping by reverse-phase HPLC and electrospray mass spectrometry (ESI-MS) were used in the characterisation of milk proteins purified from 109 individual dairy cows.

Three different PAGE systems were used. Alkaline-urea PAGE enabled the detection of α_{S1}-casein variants B and C, β-casein variants group A (variants A^1, A^2 and A^3) and B, and κ-casein variants A and B in the study population. Beta-casein variants A^1, A^2 and A^3 were subsequently resolved in an acid-urea PAGE system. The whey proteins were very poorly resolved in PAGE systems containing urea. Alpha-lactalbumin A, and β-lactoglobulin

(β-LG) variants A and B were resolved in a non-denaturing 'native' PAGE system. The frequencies of the various milk protein variants corresponded closely to figures previously published.

A free zone CE method that is able to resolve β-LG variants A, B and C was used to check the phenotype of purified β-LG samples. Three samples previously typed as β-LG BB were subsequently determined to be β-LG CC; one sample typed as β-LG BB was re-assigned as β-LG BC. This highlighted the limitations of PAGE systems for the detection of known variants.

Tryptic hydrolysis of purified casein proteins and β-LG, followed by reverse-phase HPLC separation of the resultant peptides was used to create peptide 'maps' of the hydrolysis products. Differences in peptide maps were noted between protein variants. The differences corresponded to peptides containing a substitution site. All samples analysed in this way contained more peptide peaks than expected. Analysis revealed that some were the result of incomplete digestion; others the result of chymotryptic-like cleavages. No aberrant peptide maps, indicative of a silent mutation, were detected.

Purified casein proteins and β-LG were subjected to ESI-MS for mass analysis. The mass of each protein species was determined as follows:

Protein	Average mass	Std. dev.
α_{S1}-CN B-8P	23614.9 Da	1.2 Da
α_{S2}-CN A-11P	25228.9 Da	1.5 Da
β-CN A^1-5P	24023.9 Da	3.1 Da
β-CN A^2-5P	23983.5 Da	1.8 Da
β-CN B-5P	24092.6 Da	n.d.
κ-CN A-1P	19038.8 Da	1.5 Da
κ-CN B-1P	19003.8 Da	n.d.
β-LG A	18362.6 Da	1.0 Da
β-LG B	18277.0 Da	0.9 Da
β-LG C	18287.2 Da	0.6 Da

In all cases the experimentally determined mass corresponded to the mass calculated from published primary sequences of milk protein variants.

In addition to the expected β-LG variant in each mass spectrum, additional species were detected differing from the mass of the β-LG species by increments of approximately 324 Da. Although less pronounced, the +324 Da molecular weight species were also detected in a sample of β-LG purchased from the Sigma Chemical Company. The additional species were also detected in whey prepared by ultra-centrifugation, although at a much lower level.

The 324 Da molecular weight adducts observed in ESI-MS spectra of purified β-LG are consistent with an addition of a lactosyl residue to the protein. The observation that these species remain after heat denaturation, reduction and RP-HPLC treatment suggest that the linkage is covalent. Lactulosyl-lysine is known to form in milk products during some processing conditions, particularly during heating. The observation of these glycated species in gently treated, unheated milk suggests that glycation may occur to some extent in the udder of the cow.

The association of the 324 Da molecule with β-LG does not alter the charge, molecular weight or hydrophobicity sufficiently to be detected by PAGE, CE or RP-HPLC.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the following:

New Zealand Dairy Research Institute for the opportunity to pursue a higher degree, particularly Dr Mike Boland for his guidance in the direction of this course of study.

Foundation for Research, Science and Technology and the New Zealand Dairy Board for funding.

Drs Chris Moore (senior lecturer, Biochemistry Department, Massey University) and Jeremy Hill (Section Manager, Food Science Section, New Zealand Dairy Research Institute) for their patience, tolerance and supervision during this course of study.

Jo Mudford and Carole Flyger of the Biochemistry Department, Massey University for peptide sequencing.

Professor Brian Nicholson of the Chemistry Department, Waikato University for arranging access to the department's mass spectroscopic facilities.

Wendy Jackson of the Chemistry Department, Waikato University for electrospray - mass spectrometric analysis of peptides and proteins. Particularly for her invaluable help with data analysis and instructions on how to operate the mass spectrometer.

John Allen and Daryl Rowan of the Mass Spectrometry Unit of AgResearch, Palmerston North for providing FAB-MS analysis.

Professor André Ng-Kwai-Hang of the Animal Science Department, McGill University for his encouragement, experience and enthusiasm in establishing the foundations of this work.

All of my colleagues at NZDRI, particularly Christina Coker, Wayne Thresher and Lawrie Creamer for their unflagging support, encouragement and many useful discussions on aspects of this work .

Lastly, and most importantly I wish to thank my family for their tolerance and forgiveness for many missed good night stories.

TABLE OF CONTENTS

FIGURES

TABLES

ABBREVIATIONS AND TERMINOLOGY

α-La	Alpha lactalbumin
α_{s1}-CN	Alpha S1 casein
α_{s2}-CN	Alpha S2 casein
β-CN	Beta casein
β-LG	Beta-lactoglobulin
κ-CN	Kappa casein
2-D	Two dimensional
BSA	Bovine serum albumin
CE	Capillary electrophoresis
CHO	Carbohydrate
Da	Dalton
DAD	Diode array detector
DNA	Deoxyribose nucleic acid
ESI-MS	Electrospray ionisation - mass spectrometry
FAB-MS	Fast-atom bombardment mass spectrometry
HPLC	High performance liquid chromatography
IEF	Isoelectric focusing
IEX	Ion exchange
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
RP	Reversed phase
UHT	Ultra-high temperature

In the context of this work the following terms are defined as:

Genotype - the genetic constitution of an individual organism. The genetic make up of an animal which codes for the synthesis of an individual milk protein.

Phenotype - the characteristics of an expressed milk protein as a result of the interaction of its genotype with its environment. Post-translational modifications such as phosphorylation and glycosylation frequently alter the gene product (protein) prior to expression in milk.

Electrophoretic variant - a protein variant that is able to be resolved from other genetic variants by electrophoretic techniques.

Silent variant - a protein variant with (an) amino acid substitution(s) that has no net effect on the overall electrical charge of the protein - generally undetected by standard electrophoretic techniques. Silent variants may occur when a neutral amino acid residue is substituted by another neutral residue, or when a residue carrying a charge is substituted by a similarly charged residue.